

## **Remarks**

Claims 1-24 are pending. Claims 1-22 and 24 are rejected. Claim 23 is allowable. All rejections are traversed.

Claimed is a system for encoding a plurality of videos of a moving object in a scene concurrently acquired by a plurality of cameras to generate a 3D bitstream. Camera calibration data of each camera are first determined. Then, the cameras acquire concurrently the videos. Each camera acquires one video. The camera calibration data of each camera are associated with the corresponding video. A segmentation mask for each frame of each video is determined. The segmentation mask identifies only foreground pixels in the frame associated with the object. A shape encoder encodes the segmentation masks, a position encoder encodes a position of each pixel, and a color encoder encodes a color of each pixel. The encoded data is combined into a single 3D bitstream and transferred to a decoder. At the decoder, the bitstream is decoded to an output video having an arbitrary user selected viewpoint. A dynamic 3D point model defines a geometry of the moving object. Splat sizes and surface normals used during the rendering can be explicitly determined by the encoder, or explicitly by the decoder.

Claims 1-18, 20 and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Panusopone et al., U.S. Patent No. 6,483,874 (Panusopone) in view of Carlbom et al., U.S. Patent No. 7,203,693 (Carlbom).

The Examiner states that Panusopone discloses a system for encoding a plurality of videos acquired of a moving object in a scene by a plurality of cameras. Interestingly enough, the word “camera” does not appear in Panusopone at all. The Figures of Panusopone also lack a camera. In other words, Panusopone fails to disclose any cameras. Panusopone cannot make the invention obvious. Panusopone begins with a video, and where the video came from and how it is acquired is unknown. Applicants cannot determine where the Examiner finds a camera anywhere in Panusopone. The existence of any cameras in Panusopone is, with all due respect, pure conjecture on the part of the Examiner.

As best as can be determined, the frames 105 in Panusopone are from a single video. The output of Panusopone is a 2D video object plane (VOP) on a channel 145: “The coded VOP data is then combined at a multiplexer (MUX) 140 for transmission over a channel 145.”

The only description of the Panusopone object is that it has an arbitrary shape. Panusopone does not disclose a moving object. Panusopone does not disclose determining camera calibration data of each camera of a plurality of cameras, and means for associating the camera calibration data of each camera with the video acquired by the camera. Panusopone does not disclose a segmentation mask for each frame of each video, the segmentation mask identifying only pixels in the frame associated with the moving object.

Panusopone only discloses motion vectors 220 of macro blocks. A motion vector of a macro block is not a 3D position of a pixel. With all due respect,

the Examiner has confused “motion vector” with “position,” and a “macroblock of 16×16 pixels” with a “single pixel”. The statements:

binary and gray scale shape information is encoded. With motion coding, the shape information is coded using motion estimation within a frame. With texture coding, a spatial 50

at column 5 says nothing about positions of pixels.

Those of ordinary skill in the art would know that a vector has **orientation and magnitude** while a position has **3D (x, y, z) coordinates**.

There is nothing at column 5 that discloses encoding a color of each pixel. Panusopone encodes DCT coefficients of macroblocks not individual pixel colors:

estimation within a frame. With texture coding, a spatial 50 transformation such as the DCT is performed to obtain transform coefficients which can be variable-length coded for compression.

Therefore, Panusopone cannot combine encoded segmentations masks, pixels and colors of the pixels to form a 3D bitstream encoding the plurality of videos. It would appear that Panusopone does not disclose a single one of the claimed limitations.

Panusopone does not disclose camera calibration. Carlbom discloses camera calibration. However, it is impossible to combine the teachings of Carlbom with Panusopone because there are no cameras that could be calibrated in Panusopone. Furthermore, the calibration parameters in Carlbom are not

associated with a video. Instead, the parameters are associated with an environment model:

tion parameters of cameras in the environment. Each camera has a unique identifier (ID) and its calibration parameters include its 3D position, orientation, zoom, focus, and view-  
60 ing volume. These parameters map to the 3D environment model 250, as illustrated for camera 254 in FIG. 2B.

Panusopone does not disclose any of the claimed limitations, and Carlbom cannot be combined with Panusopone. Even if it could, Carlbom fails to cure the defects of Panusopone. Thus, the invention is not made obvious.

With respect to claim 2, the arguments above apply.

With respect to claims 3-5 and 12-14, disclosed is a method for rendering from arbitrary viewpoints. This is possible because the bitstream is 3D. Neither Panusopone nor Carlbom disclose 3D rendering. The arbitrary viewpoint limitations are not addressed by the examiner.

The Examiner rejects all of the limitations in the above 8 claims with conclusory statements. As recognized in MPEP 707.07(d), “omnibus rejection of the claim ...is usually not informative and should therefore be avoided.” MPEP 707.07(f) further mandates that “where a major technical rejection is proper, it should be stated with a full development of the reasons rather than by a mere conclusion coupled with some stereotyped expression.”

The rejections by the Examiner are mere conclusions without a full development of reasons. MPEP 706.07 further makes clear that “the invention as disclosed and claimed should be thoroughly searched in the first action and the references should be fully applied.” In the present application, the rejection fails not only to provide a reasonable rationale as to how, in the examiner’s view, the applied art can be construed to teach each and every feature in the rejected claims, but the rejection also fails to even consider explicitly claimed features of the invention as recited in claims 3-5 and 12-14, and in which the entire scene is encoded using a scene specifying relations between static and dynamic portions of the scene.

The mapped trajectory in the 3D model is then related to one or more sensors within whose viewing volume the trajectory lies, as shown in FIG. 2B for the player trajectory. This is used, for example, to access video from a particular camera which best views a particular trajectory. The temporal extent of a trajectory also aids in indexing a video clip corresponding to the trajectory. As shown in FIG. 2B, the player trajectory data starting at 10:53:51 to 10:54:38 is used to index to the corresponding video clip (table 262) from the broadcast video.

As illustrated in this example, the IIMD system cross-indexes disparate data as it arrives in the database. For example, the score for a point with ID 101 is automatically related to the corresponding trajectories of the players and the ball, the exact broadcast video clip for point 101, the location of the trajectories of the players and the ball in the 3D world model, and the location, orientation and other parameters of the sensor which best views a player trajectory for the point 101. With the ability to automatically index the relevant video clips, the IIMD is also capable of storing just the relevant video alone while discarding the rest of the video data.

With respect to claim 6, see above. The Examiner has completely ignored the limitations of claim 11. The rejection of claim 11 is improper.

With respect to claim 7, as stated above, Carlbom cannot be combined with Panusopone.

With respect to claims 8, 10, 16 and 17, the Examiner does not address all of the limitations of these claims. The Examiner's rejection is an improper omnibus rejection. There is no dynamic 3D point model in Panusopone, in which the encoded segmentation masks are compressed using a lossless compression, and the position and the colors are encoded using a lossy compression, and in which the segmentation masks are encoded using MPEG-4 lossless binary shape encoding, the positions include depth values encoded as quantized pixel luminance values, and the colors are encoded using MPEG-4 video object coding, and in which the lossy compression scheme is a progressive encoding using embedded zerotree wavelet coding, and in which the shape encoder uses MPEG-4 lossless binary shape encoding, the position encoder encodes depth values, and the color encoder uses MPEG-4 video object coding.

The Examiner's rejection in rejecting all of the above limitations merely concludes:

Regarding claims 8, 10, 16 and 17, Panusopone discloses in which the segmentation masks are encoded using MPEG-4 lossless binary shape encoding, the positions include depth values encoded as quantized pixel luminance values, and the colors are encoded using MPEG-4 video object coding (col.4, ln.58 to col.5, ln.12).

With respect to claims 9-15, see arguments above.

With respect to claims 18, and 20-22, see above.

With respect to claim 19, as stated before, the teachings of Panusopone cannot be combined with Carlbom (or Wu).

With respect to claims 22 and 24, the above arguments hold. Furthermore, Rusinkiewicz does not disclose a surface normal encoder configured to encode a surface normal of each pixel, and a splat size encoder configured to encode a splat size for each pixel, and means for combining the outputs of the surface normal encoder and the splat size encoder with the single bitstream, in which splat sizes and surface normals are estimated from the positions. The Examiner does not address the limitations in these claims. The Examiner's rejection is an improper omnibus rejection:

encoder. However, Rusinkiewicz teaches the use of splat size encoder (page 344, section 2.1 Rendering Algorithms, the Qsplat uses a recursive frame rate encoding scheme for splat size encoding). Therefore, it would have been obvious to one of

It is believed that this application is now in condition for allowance. A notice to this effect is respectfully requested. Should further questions arise concerning this application, the Examiner is invited to call Applicants' attorney at the number listed below. Please charge any shortage in fees due in connection with the filing of this paper to Deposit Account 50-0749.

Respectfully submitted,  
Mitsubishi Electric Research Laboratories, Inc.

By

/Dirk Brinkman/

Dirk Brinkman  
Attorney for the Assignee  
Reg. No. 35,460

201 Broadway, 8<sup>th</sup> Floor  
Cambridge, MA 02139  
Telephone: (617) 621-7517  
Customer No. 022199